Examination of Cosmic Ray Legacy Data at Carnegie Institution of Washington.

Background. Following correspondence between Mr. S. Hardy (Carnegie Institution), the National Geophysical Data Center, NOAA, and Dr. M.A.Shea, it was recognized that an expert opinion was desirable regarding the value of the substantial amount of "legacy data" that had been stored at the Carnegie Institution following the death of Dr. S. E. Forbush. Having built and operated an ionization chamber similar to those employed by Forbush, and being aware of the various forms of errors and artifacts that arise in such data, the author was asked to examine the stored data. I was asked to provide an assessment of the value of the material, and if appropriate, provide recommendations regarding its preservation.

I spent the day of 9 July, 2004, at the Carnegie Institution. Mr. Shaun Hardy had prepared a catalogue of the contents of the 32 storage boxes, and this was used to identify ten boxes that I believed would provide me with a good insight into the nature of the contents. Each box was unpacked totally, and representative items were then examined in detail to reach the conclusions outlined below.

Analysis of Content

(1) The author has a comprehensive knowledge of all the papers written by Dr. Forbush, and it soon became clear that much of the material in the archive was of the nature of the "laboratory notebooks" that a scientist uses in the analysis of his data, and in the preparation of papers for publication. The original "scientists version" of the figures that later appeared in the Physical Review and the Journal of Geophysical Research were clearly recognizable. While interesting to examine, Forbush published well, and there was nothing (in the sample seen) that would warrant preservation of such material except for historical purposes.

(2) The catalogue referred to "cosmic ray traces", and these were examined with care. They were the original photographic strips on which each days data were recorded by the automatic recording system on the ionization chambers. There was extensive data from all four ionization chambers. Each daily strip was approximately 2x 24 inches in dimension; each month of data was carefully stored between robust cardboard or composition board strips; and the time annotation was faint but adequate. (Note however later comment). Some of this data was examined in detail. It was in remarkably good condition. It had clearly been stored in a dry place- there was absolutely no adhesion of the sheets. The photographic material was in excellent condition- there was no sign of poor fixing, or fading. The contrast on most strips was excellent, and would permit machine based reading. The trace for the 25July 1946 "solar flare event" was examined carefully. It was in remarkably good condition. I have been a contributor to the literature in the analysis of the solar flare effect and I examined this trace with the question in mind "Is there anything about this record that has not been captured in Forbush's several papers, or that, now with the hindsight of 50 more years, may make it desirable to conduct further analysis.". My answer was "yes, there is new insight here". See recommendations.

(3) Forbush published the bihourly data from four of the ionization chambers for the period starting in 1936/7- 1965????, and these are an adequate record for most purposes. The question then was: "is there material in the archive that could usefully extend that existing legacy record". The answers were;

(3A)There are data in the archive prior to the commencement of the published record. Data were sighted from Cheltenham for 1935 and 1936, and from Chicago (where the instruments were built) from 1934. The cosmic ray variations between the solar minimum of 1934, and the commencement of the published tabulations in 1936/7 are of considerable interest, and these data could be of considerable interest. Note however the warning in paragraph 4.

(3B) There are tabulations of the hourly pressure corrected hourly data that would be valuable for investigations of selected periods, such as the commencements of large "Forbush decreases".

(3C) There is a considerable quantity of the data (starting 1937) from the ionization chamber in Mexico City in the archive (not much of this was included in the published tabulations). See however the warning in paragraph 4.

(4) The ionization chamber was an idiosyncratic instrument. There were many factors that corrupted the data that were known only to the custodians, or the custodians of similar instruments. The most serious were (a) long term drift due to changes in radioactive contamination; (b) step changes due to changes in the instrument configuration (several being noted in my inspection); (c) deliberate, or inadvertent changes in the "balance current", (d) pressure leaks (records of two seen in my brief inspection); and (e) changes in the sensitivity of the recording electroscopes. There were many others. I saw examples of problems that were never recorded in the literature. The original day-to day "log books" are included in the archive for each instrument, and some were examined. However, Forbush was a careful worker, and his published papers record his efforts (and dissatisfaction with the result) to eliminate these changes over the long term. I take the view that he will have done as good a job as possible, and that it would be dangerous to retain data that, in inexperienced hands, could be at variance to his carefully corrected data.

(5) To my surprise, I failed to find any record of the "time calibrations". An ionization chamber recording system was connected to a clock that earthed the central collector each hour, and this clearly defines a time scale that can be (with care) read to the minute on the traces I examined. However, it was then necessary to record the (time dependent) differences between the "station clock" and Universal Time (eg, using WWV). This was normally done on a weekly basis. One of the greatest values of the legacy data will be the ability to study the cosmic ray flare effects on the records in greater detail. Such studies will be impeded if the time calibrations are not there (surprisingly, they were not on the station log I

examined). However, it was clear from the archive (and I already knew) that Forbush was almost obsessive about his record keeping, and therefore I expect that the data exists in the archive. They may require careful search, and may be difficult tom recognize.

Recommendations.

The cosmic ray record as recorded by instruments, and in the cosmogenic isopes, shows that the character of the long term changes in the cosmic ray flux changed significantly about 1954; prior to that, there had been a long term steady decline in the cosmic ray intensity from 1900 ; after 1954, this decline had stopped. Solar activity had been steadily increasing during this interval as well to its high point in 1957. While the world wide neutron monitor network provides excellent detail of the temporal behaviour from the International Geophysical Year, onwards, it is very sparse in the interval 1951-56, and non- existent prior to that. That is, the Forbush record is the only comprehensive record of the period of long term decline in the cosmic radiation prior to 1954. This is the primary reason for maintaining the Forbush data as an important part of our scientific legacy. Its may have its greatest value when the solar activity commences declining (probably in the relatively near future) to the low values last seen circa 1890. On that basis, I provide the following recommendations.

(A) The bihourly data compilations published by Forbush and his coworkers are undoubtedly a major component of his legacy data. I presume that all four volumes are already incorporated in the national data base in digital form. If not, they should be.

(B)As noted above, there may be data from Mexico City that were not included in his published tabulations (I saw evidence, new to me, that he may not have had control of that instrument). If so, I recommend that the Mexican bihourly data in the archive be collated, examined, and then added to the existing set of legacy bihourly data in digital form. (Expert assessment necessary).

(C)That the unpublished data in the archive from 1934, 35, and 36 be examined with care, and if appropriate, converted to the bihourly form to extend the existing legacy data back to the sunspot minimum of 1934. (Expert assessment mandatory).

(D)With the benefit of the past 70 years of progress, and future progress, it is likely that the short term (~ minute) time detail in the solar flare effects will provide new understanding. Also, modern signal processing methods may allow smaller flare effects to be seen by virtue of the inherent 1 minute time resolution of the record. (This would be particularly valuable for the period prior to the International Geophysical Year of 1957, from which time there is good world wide coverage using the more sensitive neutron monitor network). It is therefore recommended in the strongest terms that the original photographic records be digitized on a minute by minute basis. This would imply a data record of 500,000 bytes per year, per station. As noted above, it appears highly likely that this could be automated to a high degree. (Expert assessment of the time calibration necessary, and in drawing up the data reqaqding specification).

(E)That the hourly data in the archive for large Forbush decreases in the period 1935-1957 be copied into a digital form, and added to the existing bihourly legacy data.

(F)That the majority of the material in the archive, being of the form of Forbush's personal research notes, can be disposed of following careful examination to identify those items of historical significance. (Expert assessment necessary).

Operational Considerations.

It is stressed that my examination was limited to a small subset of the total archive. I saw much that has not been recorded in the literature. I therefore recommend the following step by step process.

(1) That the methodology for the digitization of the original photographic record be examined. To this end, but to preserve the integrity of the archive, it is recommended that a single month of the Cheltenham data after 1957 be extracted from the archive and used for evaluation and test purposes (chosen to have good contrast). The author proposes that the goal should be to achieve one sample per minute. The time calibration of the record will require careful consideration.

(2) That the existing archive be sorted in a comprehensive manner, with careful attention to both the historical and legacy data aspects of the archive, as summarized above. It is estimated that this would reduce the archive to 25% of its present size. It is stressed; someone with knowledge of the ionization chamber technology should control this process. Ten days would be required to do it justice.

(3) That the appropriate institutional and contractual arrangements be put in place to implement the recommendations (A) through (E) above.

K.G.McCracken
Maryland,
11 July, 2004.